AIDS, Publish Ahead of Print

DOI: 10.1097/QAD.0000000000001778

No Neurocognitive Advantage forImmediate Antiretroviral Treatment inadults with greater than 500 CD4+ T Cell Counts

Edwina J Wright, Birgit Grund, KevinR. Robertson, LucetteCysique, Bruce J Brew, Gary Collins, Mollie Poehlman-Roediger, Michael J Vjecha, AugustoPenalva de Oliveira, Barbara Standridge, Cate Carey,AnchaleeAvhingsanon, Eric Florence, Jens D. Lundgren,Alejandro Arenas-Pinto, Nicolas J Mueller, Alan Winston, Moses Supercharger Nsubuga, Luxshimi Lal, Richard W. Pricefor the INSIGHT START Neurology Substudy Group

Affiliations

Edwina J. Wright

Department of Infectious Diseases Alfred Health, Monash University, Burnet Institute, The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

B**irgit Grund:**

School of Statistics, University of Minnesota, Minneapolis, MN, USA

Kevin R. Robertson

Department of Neurology, University of NC, USA

LucetteCysique

Neurosciences Research Australia, The University of New South Wales,
St. Vincent's Hospital Sydney, Applied Medical Research Centre, Peter Duncan Neuroscience Unit,

Sydney Australia

Bruce J. Brew

University of New South Wales, Neurosciences Program and Peter Duncan Neurosciences Unit

St Vincent's Centre for Applied Medical Research,Sydney, Australia

Gary Collins

Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA

Mollie Poehlman-Roediger

Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA

Michael Vjecha

Institute for Clinical Research, Inc., Washington, DC

Augusto Penalva de Oliveira,

Instituto Emilio Ribas, Sao Paulo, Brazil

Barbara Standridge

Veterans Affairs Medical Center, Washington DC, USA

Cate Carey

Kirby Institute, University of New South Wales, Sydney, Australia

AnchaleeAvihingsanon, MD, PhD

HIV-NAT Thai Red Cross AIDS Research Centre and Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Eric Florence

Department of Clinical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

Jens D. Lundgren

CHIP, Department of Infectious Diseases, Rigshospitalet, University of Copenhagen, Denmark

Alejandro Arenas-Pinto MBBS, MSc, PhD

MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology, University College London

Nicolas J. Mueller

Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich

Zürich, Switzerland

Alan Winston

Division of Infectious Diseases, Department of Medicine, Imperial College London, London W2 1NY

Moses Supercharger Nsubuga,

People in Need Agency (PINA), Uganda

Luxi Lal

Burnet Institute, Melbourne, Australia

Richard W. Price

Department of Neurology, University of California San Francisco, San Francisco, CA, USA

Dedication to Mollie Poehlman-Roediger

This paper is dedicated to the memory of Mollie Poehlman-Roediger, our kind and joyful colleague who worked with us closely on this project and who is greatly missed. Vale.

Running Head: START Neurology Substudy

Abstract

Objective: To compare the effect of immediate versus deferred antiretroviral treatment (ART) on neuropsychological test performance in treatment-naive HIV-positive adults with >500 CD4+ cells/μL.

Design: Randomized trial.

Methods: The START parent study randomized participants to commence immediate versus deferred ART until CD4+ <350 cells/μL. The START Neurology substudy used 8 neuropsychological tests, at baseline, months 4, 8, 12 and annually, to compare groups for changes in test performance. Test results were internally standardized to z-scores. The primary outcome was the average of the eight test z-scores (QNPZ-8). Mean changes in QNPZ-8 from baseline were compared by intent-to-treat using longitudinal mixed models. Changes from baseline to specific time points were comparedusing ANCOVA models.

Results:592 participants had a median age of 34 years; median baseline CD4+ count of 629 cells/μL; the mean follow-up was 3.4 years. ART was used for 94% and 32% of accrued person-years in the immediate and deferred groups, respectively. There was no difference between the immediate and deferred ART groups in QNPZ-8 change through follow-up [-0.018 (95% CI: -0.062 to 0.027, p=0.44)], or at any visit. However, QNPZ-8 scores increased in both arms during the first year, by 0.22 and 0.24, respectively (p<0.001 for increase from baseline).

Conclusions: We observed substantial improvement in neurocognitive test performance during the first year in both study arms, underlining the importance of using a control group in studies assessing neurocognitive performance over time. Immediate ART neither benefitted nor harmed neurocognitive performance in individuals with CD4+ cell counts above 500 cells/μL.

Keywords: HIV, Neurocognitive impairment, Antiretroviral Treatment, Central Nervous System, HAND

Research in Context

Evidence before this study

In advanced, untreated Human Immunodeficiency Virus type 1(HIV-1) infection, 15%- 20% of individuals develop HIV-associated dementia (HAD). Combination antiretroviral therapy (ART) improves neuropsychological performance in 40%-60% of individuals with HAD. In acute HIV-1 infection, mild neurological manifestations occur in up to 50% of individuals and clinical resolution is usually observed with immediate ART. However, it is unclear if such benefit occurs in HIV-positive individuals with high CD4+ cell counts and whether any benefit might be counteracted by potential ART toxicities, which have been reported in individuals receiving ART regimens that have high penetration into the brain.

Added value of this study

We undertook a Neurology substudy within the Strategic Timing of Antiretroviral Treatment (START) study. The START Neurology substudy which enrolled 608 participants is the largest clinical trial to date to evaluate the impact of ART on neurocognitive performance in HIV-positive, ART-naïve individuals with > 500 cells/μL. Our study found that participants randomised to commence ART immediately versus deferring ART until CD4+ < 350 cells/μL did not experience either benefit or harm with respect to their neurocognitive performance, during a mean follow-up period of 3.4 years. ART regimens with high brain penetration did not benefit or advantage either treatment group. Importantly we observed a marked improvement in neurocognitive test performance in both study arms during the first 12 months, strongly suggesting a practice effect.

Implications of all the available evidence

Our finding suggests that there was minimal underlying neurological damage that could be either prevented or reversed by immediate ART in this study population. The START parent study showed that immediate versus deferred ART decreases the risk of serious AIDS and non-AIDS

illnesses by 57%. These pivotal findings led to the 2015 World Health Organisation recommendation that all HIV-positive individuals should initiate ART irrespective of CD4+ cell counts, and the START Neurology substudy findings support the safety of initiating ART with respect to neurocognitive performance. Our study also underlines the importance of having a control arm in intervention studies that evaluate neurocognitive test performance over time.

INTRODUCTION

In advanced, untreated HIV infection, 15%- 20% of individuals develop HIV-associated dementia (HAD)[1,2], the severe form of HIV-associated neurocognitive disorders (HAND)[3].HAND is a subcortical dementia that results in psychomotor slowing and is associated with increased risk of mortality[4], job loss[5] and poor medication adherence[6].Combinationantiretroviral therapy (ART) improvesneuropsychological performancein40%-60% of individuals with HAD[7-9]. ART regimens with higher versus lower CNS penetration may effect greater improvement in individuals with HAD[10,11].

In acute Human Immunodeficiency Virus type 1 (HIV-1) infection, mild neurological manifestations occur in up to 50% of individuals and prompt clinical resolution is usually observed with immediate ART[12]. In this setting, neuronal injury may occurwith raised levels of neurofilament light chain in the cerebrospinal fluid (CSF) and altered ratios of CNS metabolites in brain magnetic resonance imaging[13]. Therefore, plausibly, early ARTmaypreserveneurological function, or reverse neurological damage caused byHIV infection. However, it is unclear if such benefitoccurs at high CD4+ cell counts and whether it might be counteracted by potential ART toxicities[14].

We undertook the Neurology substudy of the Strategic Timing of Antiretroviral Treatment (START) trialto test the hypothesis that immediate versus deferredART would benefitneurocognitive performancein antiretroviral-naïve adults with > 500 CD4+ cells/μL.

METHODS

Study design

START is a large international, multicentre clinical trial, performed by the International Network for Strategic Initiatives in Global HIV Trials (INSIGHT). START randomized 4,684 ART-naïve, HIV+ participants with CD4+ counts > 500 cells/μL to receive immediate versus deferredART untilthe CD4+ cell count fell to < 350 cells/μL[15].At selected sites, the Neurology Substudyco-enrolled participants who underwenta standard neuropsychological test battery at baseline, months 4, 8, 12, and annually thereafter to compare changes in neuropsychological test performance in the immediate versus deferred groups, (described elsewhere).[16]We report

resultson data accrued through May 26, 2015, the day beforethe parent START study wasunblinded and all participants in the deferred ART group wererecommended to initiate ART because immediate ART was found to have decreased the risk of serious AIDS and non-AIDS illnesses by 57%[15].

Study population

The START Neurology substudy co-enrolled 608 participants between May 2009 and June 2012 at 35 sites in Argentina, Australia, Belgium, Brazil, Chile, Germany, Italy, Switzerland, Thailand, the United Kingdom, and the USA. At participating sites, all eligible subjects were offered substudyco-enrolment. Eligibility criteria included STARTco-enrolment and ability to perform the studytests.The substudy was approved by each institution's Institutional Review Board.Participant information and consent forms were translated as required.All participants provided written informed consent.

Neuropsychological test battery

The neuropsychological test battery consisted of 8 tests (grooved pegboard, finger tapping, Color Trails 1 and 2, Semantic Verbal Fluency, WAIS-III Digit Symbol, HVLT-R Learning, HVLT-R Delayed Recall), coveringsixcognitive domains (Supplemental Appendix, Table S1, http://links.lww.com/QAD/B233 footnote) that are affected by HAND[3]. The test battery was constructed to be adaptable across different cross-cultural, international settings[17],to be brief, easy to administer andscore, and sensitive to HIV-associated brain injury[3].

The Center for Epidemiological Studies-Depression (CES-D) scale[18]was administered at each substudy visit. A CES-D score ≥16 was considered to indicate depression.

Details of training, translations, administration of tests and questionnaires, and staff accreditation are described elsewhere[16].

Outcome measures

Test scoreswere standardized to z-scores using thebaseline test results of the 608 study participants as a reference (internal standardization), resulting in mean=0, SD=1 at baseline for each test (for detailed description and rationale see[16]). Standardization of the Semantic Fluency

test was an exception:because we used alternate Semantic Fluency test versions across visits to minimize practice effect,we standardized its z-scores at each follow-up visit to mean=0 and SD=1 using each visit'spooled study population as reference. The quantitative neuropsychological performance z-score (QNPZ-8) was calculated as the average of the z-scores for the 8-test battery.

Primary outcome

Theprimary outcome was the change in theQNPZ-8 from baselinethrough follow-up.

Secondary outcomes

Secondary outcomes included individual test z-scores, neurocognitive impairment (NCI), and a Global Deficit Score (GDS). We defined NCI based on a cognitive domain impairment score rating in line with the Frascati Criteria[3]; we did not assess for functional status or confounding factors.Mild NCI (comparable to asymptomatic neurocognitive impairment and early mild neurocognitive disorder [MND][3])was defined as having internally standardized z-scores −1 SD below thesample mean of zero in two or more of the six cognitivedomains; moderate/severe NCI (comparable to advanced MND and HAD[3]) was defined as having z-scores −2 SD below thesample mean of zero in two or more domains.

The GDSwas computed as the average of deficit scores for the 8 tests; the deficit scores grade normal performance and impairment into 6 categories: 0 (normal), $z \geq -1.0$; 1 (mild), $-1.0 > z \geq -1.5$; 2 (mild-moderate), $-1.5 > z \geq -2.0$;3 (moderate), $-2.0 > z \geq -2.5$; 4 (moderate-severe), $-2.5 > z \geq -3.0$; and 5 (severe), $z < -3.0$. With the GDS,low performanceon some tests is not cancelled out by high performance on others[19].

ART use

ART regimens were selected ("pre-specified") prior to randomization by site investigators from a table of regimens recommended by the US Department of Health and Human Services (see supplementary appendix in[15], http://links.lww.com/QAD/B233). We calculated CNS penetration efficiency (CPE) scores of participants' ART regimens using the 2010 version[20].

Other data collection

In addition to data collectedin the parent study[15], the Neurology substudycollected rural or urban residence, current employment status and education level. The Framingham 10-year risk of coronary heart disease was calculated as a cardiovascular health measure[21].

Statistical Methods

The primary analysis for the Neurology substudywas an intent-to-treat comparison between the immediate and deferred ART groups for changes in QNPZ-8 from baseline through follow-up, using a longitudinal mixed model with an indicator variable for treatment group, adjusted for visit and for baseline QNPZ-8 scores.The sample size of 600 participants was estimated to detect an average treatment differencein the change in QNPZ-8 scores between the two study arms of 0.13with 80% power at a 5% significance level.

A detailed description of statistical methods was included in the Supplemental Appendix, http://links.lww.com/QAD/B233. For all analyses, follow-up was censored at each participant's last visit prior to May 27, 2015, when the parent START study wasunblinded. To illustrate the differential use of ART in the immediate and deferred groups and its effect on CD4+ cell counts and HIV RNA levels through follow-up, the proportion of participants using ART, the proportion of participants with HIV RNA $\leq$ 200 copies/mL and the mean change in CD4+ cell counts were summarized by treatment group in 4-month intervals.The treatment difference in change in CD4+ cell counts through follow-up was estimated in a longitudinal mixed model adjusted for visit and baseline CD4+.

In addition to the primary analysis, we also compared treatment groups for changes in QNPZ-8 from baseline through month 12 only. By design, participants in the immediate group were to initiate ART at randomization, while few participants in the deferred group initiated ART within the first year; therefore, the difference between treatment groups over the first 12 monthsis an approximate estimate of the effect of ART versus no ART use. We performed similar intent-to-treat comparisons of changes in z-scores for each of the 8 tests as planned per protocol. Within each treatment group, changes in QNPZ-8 and individual z-scores from baseline to annual visits were summarized by means with 95% confidence intervals (CIs). Groups were compared for changes from baseline to each visit using t-tests in linear regression models adjusted for baseline

scores. We compared treatment groups for changes in the prevalence of NCI and depression from baseline through follow-up using generalized estimating equations (GEE) for binary responses, and used Chi-squared tests to compare prevalence at each visit. We compared treatment groups for changes in GDS using longitudinal mixed models adjusted for visit and baseline GDS. We compared treatment groups for changes in CES-D scores using similar longitudinal mixed models.

To assess the effect of ART versus strictly untreated HIV, we also compared the immediate group (excluding participants who did not start ART within the first year) versus the deferred group (censored at ART start) for changes in QNPZ-8 and individual z-scores; this comparison is not protected by randomization.

Subgroup analyses for the primary endpoint were performed to determine whether the treatment effect differed across baseline characteristics. The homogeneity of the treatment effect across subgroups was assessed by testing for interaction between the subgroup variable and treatment group indicator in longitudinal mixed models; when possible, the continuous subgroup variable was used to test for homogeneity.To adjust for multiple comparisons, we used the Benjamini-Hochberg method to limit the false discovery rate (FDR) to 5%.

Analyses were performed with SAS version 9.4 (SAS Institute, Cary, North Carolina, United States) and R version 3.[22] All p-values are two-sided; $p \leq 0.05$ denotes significance.

RESULTS

Baseline demographics, laboratory and clinical characteristics

Key baseline characteristicsof the 608 substudyparticipantsare summarized in Table 1;592 participantshad neuropsychological test data at baseline and follow-up, and were included in the current analyses (Fig. 1). Using the cognitive domain impairment rating, we found that19.8% of participants were at least mildly impaired, and2.7% were moderately or severelyimpaired; the median GDS was 1[IQR 0-3] (Table 1). Baseline neuropsychological test results are summarized in Table S1 (Supplemental Appendix, http://links.lww.com/QAD/B233).There was no differencebetween study arms inany of the baseline factors.

ART use, HIV RNA, and CD4+ cell counts through follow-up

Participants were followed for a mean of 3.4 years (range 0.2 – 5.4 years). By design, ART use differed substantially between treatment groups. In the immediate ART group, 291 (93.1 %) of participants started ART within 2 months of randomization, and 92.8% or more used ART at any follow-up visit (Fig.2A). In the deferred ART group, 11.1%,32.8%, 52.3% and 63.6% were using ART at months 12, 24, 36, and 48, respectively (Fig.2A). ART was used for 94.2% of follow-up time accrued in the immediate group, and for 31.8% in the deferred group (Fig.2B). During the first year, ART was used for 91.5% of follow-up time in the immediate group compared with 4.4% in the deferred group.

Differences in ART use between the groups were reflected in the HIV RNA and CD4+levels. For almost all participants, viral load was suppressed while using ART (Fig.2A). Through follow-up, mean CD4+ cell counts were higher in the immediate ART group, by 226 cells/μL (95% CI 201-250, p<0.001) (Fig. 2C). Mean CD4+ cell counts at ART commencementwere 676 and 411 cells/μL in the immediate and deferred ART groups, respectively.

Neuropsychological test performance through follow-up

The trajectories for mean change in QNPZ-8 in the immediate and deferred ART groups were almost identical;mean QNPZ-8 scores increased substantially from baseline through month 12,by 0.22 and 0.24, respectively (p<0.001each for increase), and remained stable afterwards (Fig.3A, and Supplemental Appendix, Table S2A, http://links.lww.com/QAD/B233). There was no difference between treatment groups in change in QNPZ-8 from baseline through follow-up (estimated difference -0.02 [95% CI: -0.06 - 0.03, p=0.44]), or from baseline to any of the follow-up visits (Fig.3, andSupplemental Appendix, Table S2A, http://links.lww.com/QAD/B233).

When considering individual tests, we found no difference between treatment groups for 7of the 8tests (p=0.08 to 0.94 for comparing mean change in z-scores through follow-up)(Fig.3B, and Supplemental Appendix, Tables S2B-I, http://links.lww.com/QAD/B233).For the Digit Symbol test,while performance increased in both arms, the z-score increase was lower in the immediate ART group, with an estimated treatment differencethrough follow-up of -0.12 (95% CI: -0.21to -0.04, p=0.005) favouring the deferred ART group (Supplemental Appendix, Table S2D, http://links.lww.com/QAD/B233).

The pattern of an initial marked increase in z-scores through month 12 in both treatment groups was apparent for the Grooved Pegboard,ColorTrails 1 and 2, and WAIS Digit Symbol tests. Z-scores for the HVLT-R Learning test also increased over time. For Semantic Verbal Fluency, only the treatment difference, but not the overall increase or decrease from baseline could be estimatedbecause z-scores for this test were standardized to zero at each follow-up visit to account for the different test versions used at different study visits.

Sensitivity analysesshowed similar results, when comparing treatment groups through month 12 only (Supplemental Appendix, Tables S2A-I, http://links.lww.com/QAD/B233), and when excluding participants who did not start ART within the first year in the immediate ART group and censoring follow-up at ART initiation in the deferred group (Supplemental Fig. S1, http://links.lww.com/QAD/B234).

The prevalence of NCI (not corrected for practice effect at follow-up) and changes in mean GDS are shown in the Supplemental Appendix, Tables S3 and S4, http://links.lww.com/QAD/B233, respectively; there was no evidence for a difference between treatment groups.

Depressive symptoms

There was no difference between treatment groups regarding change in continuous CES-D scores, estimated difference -0.59 (95%CI: -1.63 - 0.45, p=0.27) for longitudinal comparison(Supplemental Appendix, Table S5, http://links.lww.com/QAD/B233). Depression prevalence(CES-D $\geq$ 16) was similar in bothgroups (p=0.21) (Supplemental Appendix, Table S6, http://links.lww.com/QAD/B233).

Subgroup analyses

Fig.4 illustrates treatment differences for the change in QNPZ-8 acrossseveral subgroups; of these, subgroup analysesby age, education, baseline HIV RNA, baseline QNPZ-8, pre-specified ART regimens, and CPE scorewere defined a-priori. We analysed 24 subgroup factors, listed in the footnote to Fig.4.

Among participants whose pre-specified regimens included efavirenz,the deferred ART groupimproved their QNPZ-8 scoreslightly more than the immediate group, estimated mean difference -0.05 (95%CI: -0.10 – -0.00). In contrast, among participants with other pre-specified ART, the immediate armshowed greater improvement in their QNPZ-8 scores than the deferred arm, estimated difference 0.11 (95% CI: 0.01 – 0.20) (p=0.004 for heterogeneity of the treatment effect) (Fig. 4).Importantly, participants who were pre-specifiedART without efavirenz differed from those who were pre-specified efavirenz in several characteristics that may impact upon neurocognitive test performance, including a higher prevalence of prior psychiatric diagnoses (20.6% versus 4.9%) and depression(CES-D score ≥ 16, 43.3% versus 28.6%) (Supplemental Appendix, Table S7, http://links.lww.com/QAD/B233).

Additionally, the treatment difference between the immediate and deferred groups varied across subgroups by baseline QNPZ-8 scores (p<0.001 for heterogeneity) and by the baseline global deficit score (p=0.004). There was no evidence for a difference in mean QNPZ-8 change between the immediate and deferred ART groups within any of the investigated subgroups, however, except for the subgroup of participants who were not pre-specified EFV (Fig. 4).

DISCUSSION

The START Neurology substudyis the largest controlled clinical trial to evaluate the impact of ART on neurocognitive performance amongHIV-positive individuals with > 500 cells/μL. We found no difference in the change in neuropsychological test performance when comparing immediate versus delayed ART in previously untreated, HIV-positiveadults with CD4+ cell counts above 500 cells/μL. Thus, the study'shypothesisthat immediate versus delayed ARTwould have a favourable effect on neurocognitiveperformancewas refuted. As a corollary, we found that early ART neither benefits nor harms neurocognitive performance.

Why was there no beneficial effect of immediate ART on neurocognitive performance, given that benefit has been reported in previous studies?[7-9]This study was well-powered to detect a modest treatment difference. There was no difference between the outcomes in the two arms evaluated either by intent-to-treat or in sensitivity analyses. During the first year, ART was used for 91.5% of the follow-up time accrued in the immediate group, compared with 4.4% in the

deferred group and undetectable HIV-RNA levels were observed on ART;therefore, the study's finding could not be explained by ineffective ART, or poor adherence.

It is highly likely that practice effect influenced the sharp, near-identical increase in aggregate test performance (QNPZ-8) in both study arms through month 12. Practice effect occurs following the repeated administration of neuropsychological tests and is well-documented[23,24], but often ignored in Neuro-HIV studies[25]. The trajectories we observed aresimilar to those seen with repeated testing in healthy persons, or in HIV-positive persons who are clinically stable on ART[26]. Of note, participants in both study arms achieved the same incremental improvement in QNPZ-8 from baseline to year one, and the improvementwas orders of magnitude larger than anydifferences between the immediate and deferred ARTgroups. In previousNeuro-HIV studies that reported beneficial effects of ART, all participants started ART at study entry, there was no control group of delayed ART, no adjustment for practice effect, and the observed improvement in test performance was attributed to ART[7-9]. Our findingscontradict conclusions drawn from uncontrolled prospective studies and underline the importance of a control arm in studies assessing neurocognitive test performance over time.

The likeliest biological explanation for the observed lack of effect of immediate ART is that there was little HIV-induced neural injury in our study population, despite the presumed presenceof HIV in CSF and local inflammation within at least some of the participants. Study participants wereyoung, urban, educated, mostly employed, and without rapid immune progression;these factorsmay have afforded neuropsychological protection against the effect of HIV, and as a result neuropsychological performance was not remediated by immediate ART.

With respect to the possibility that ART may have contributed to CNStoxicity and hence abrogated any possible benefit of immediate ART, use of ART regimens with high CNS penetration effectiveness scoresneither benefitted, nor disadvantaged either treatment group. Similarly, with respect to the potential toxicity from efavirenz, those participants whose pre-specified ART regimen did not include efavirenz had slightly greaterneuropsychological improvement in the immediate versus the deferred group. However, comparing efavirenz to other ART is basedon a non-randomized analysis that needs to be interpreted cautiously, as those participants who were pre-specifiedefavirenzdiffered markedly from those pre-specified other ART.

It is possible that immediate ART in individuals with high CD4+ cell counts protects neural health in ways that were not captured by the neurocognitive tests used in this study, or that will only manifest in a delayed fashion. For example, early treatment might reduce or stop expansion of the CNS HIV reservoir, potentially resulting in longer-term benefit.[27] Other measures of ongoing neural injury, including CSF or blood biomarkers such as neurofilament light chain (NFL)[28-30] or neuroimaging modalities,[31] may eventually prove to be more sensitive and robust than the neurocognitive test performance. On the other hand, in the absence of confounding conditions, neurocognitive performance has been the evaluation and diagnostic standard for assessing the impact of HIV on CNS functional integrity,[32,33] and was not appreciably altered by early compared to delayed therapy in this study.

The study's chief strengths were itsrandomized design, the large sample size, and the standardizedtest battery administration. There were several limitations. First,mean follow-up was 3.4years and,plausibly, an ART effect couldemerge after longer treatment duration. Notwithstanding, there was no evidencefor a divergence of treatment arms in the study's later years. Second, the test battery was limited to eightneuropsychological tests. However, the battery comprises tests shown to be highly correlated with cognitive performance on a larger battery[34]. Lastly, we have not measured biomarkers of neural injury or the size of the HIV reservoir in cerebrospinal fluid.

In summary,we observed a striking improvement of test performance during the first year in both study arms, which underlines the need for a control group in studies assessing neurocognitive test performance over time. The parent START study showed that immediate ART significantly decreases risk of serious AIDS and non-AIDS conditions, leading to the 2015 WHO recommendation that all HIV-positiveindividuals should initiate ART irrespective of CD4+ cell counts[35]. However, the START Neurologysubstudyshows neitherbenefit, nor harm of early ARTwith respect to neurocognitive performance in individuals with CD4+ cell countsabove 500 cells/μL.

Funding

The parent START study was supported by the National Institute of Allergy and Infectious Diseases (United States), AgenceNationale de Recherches sur le SIDA et les HipatitesVirales

(France), National Health and Medical Research Council (Australia), National Research Foundation (Denmark), BundesministeriumfürBildung und Forschung (Germany), European AIDS Treatment Network,Medical Research Council (United Kingdom), National Institute for Health Research, National Health Service (United Kingdom), and the University of Minnesota. Antiretroviral drugs were donated to the central drug repository by AbbVie, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline/ViiV Healthcare, Janssen Scientific Affairs, and Merck. Additionally, the National Institute of Mental Health and the National Institute of Neurological Disorders and Stroke(United States) specifically funded the START Neurology substudy.

NIH Grants: UM1-AI068641 and UM1-AI120197, NINDS/NIMH (funding provided via START NIH grant).

Role of Authors:

EJW, RWP, KR, BB, and BG designed the study, EJW, BG and RWP wrote the first draft of the manuscript, BG, GC and MPR analysed the data. EJW, KR, LC, BB, MV, APdO, BS, AA, POJ, EF, JL, AAP, NM, AW, LL and RWP enrolled participants, performed training, and/or supervised implementation of the study. All co-authors critically reviewed the manuscript.

Acknowledgments

The authors wish to acknowledge and thank the study participants.

START Neurology Substudy Credit Roster

International Coordinating Centers

Copenhagen: PO Jansson.

London: AG Babiker, A Arenas-Pinto, N B-Atako, E Dennis, S Forcat, F Hudson, B Jackson, C Purvis, C Russell.

Sydney: S Emery, C Carey, M Clewett, S Jacoby.

Washington: B Standridge, A Sanchez, MJ Vjecha.

Site Investigators by Country by Enrollment

(n=number of participants enrolled)

Brazil (n=169)

Projeto Praça Onze Pesquisa em Saúde (n=102): SR Telles, NN Tebet.

Instituto de InfectologiaEmílioRibas – IIER (n=67): ACP Oliveira, MRP Gascon.

Thailand (n=89)

Chulalongkorn University Hospital (n=68): K Ruxrungtham, S Gatechompol.

KhonKaen University, Srinagarind Hospital (n=21): P Chetchotisakd, S Anunnatsiri.

Site Coordinating Center: A Avihingsanon, P Rerksirikul.

United States (n=88)

Denver Public Health (n=17): J Scott, E Gardner.

Regional Center for Infectious Disease (n=12): K Epperson, C Van Dam.

UNC AIDS Clinical Trials Unit (n=11): MR Chicurel-Bayard, D Currin.

Virginia Commonwealth University (n=11): V Watson, DE Nixon.

The R & E Group at the Portland VA Research Foundation (n=8): MD Murphy, SM Sweek.

Bronx-Lebanon Hospital Center (n=7): R Cindrich, M Vasco.

Naval Medical Center San Diego (n=7): MF Bavaro, SJ Echols, BK Agan.

San Antonio Military Health System (n=6): JF Okulicz, TJ Sjoberg.

Wayne State University (n=4): M Farrough, R MacArthur.

Wake County Human Services (n=2): C Kronk, J Jackson.

(Closed sites not included)

Belgium (n=59)

Centre HospitalierUniversitaire St. Pierre (C.H.U. St. Pierre) (n=30): K Kabeya, V Lenoir.

Institute of Tropical Medicine (n=29): M van Frankenhuijsen, L van Petersen.

United Kingdom (n=48)

St. Mary's Hospital (n=19): B Mora-Peris, A DelRosario.

Barts and the Royal London (n=12): C Orkin, J Hand.

Chelsea and Westminster Hospital, London (n=12): B Gazzard, C Higgs.

St. Thomas' Hospital (n=5): J Fox, A Sharp.

Argentina (n=46)

FUNCEI (n=22): G Lopardo, GL Copertari.

Hospital General de Agudos JM Ramos Mejia (n=15): MH Losso, J Bruguera.

Hospital Rawson (n=9): D Daniel, A Crinejo.

Site Coordinating Center: GR Loria, ML Doldan, A Moricz.

Chile (n=38)

Fundación Arriarán: M Wolff, G Allendes.

Germany (n=24)

Klinik I für Innere Medizin, Klinikum der Universität zu Köln (n=14): C Lehman, C Wyen.

Medizinische Universitätsklinik - Bonn, Immunologische Ambulanz CRS (n=6): J Rockstroh, C Schwarze-Zander.

Johann Wolfgang Goethe - University Hospital, Infektionsambulanz CRS (n=4): C Stephan, T Wolf.

Australia (n=23)

The Alfred Hospital (n=11): J Hoy, J Costa.

St Vincent's Hospital Sydney (n=10): DA Cooper, K MacRae.

Sexual Health and HIV Service - Clinic 2 (n=2): D Rowling, E Warzywoda.

Site Coordinating Center: S Emery, C Carey, M Clewett, S Jacoby.

Switzerland (n=14)

University Hospital Zurich (n=9): N Müller, M Rizo-Oberholzer.

Bern University Hospital (n=5): H Furrer, M Lacalamita.

Italy (n=10)

IRCCS San Raffaele, Milan: P Cinque, F Ferretti.

The complete list of START investigators can be found at N Engl J Med 2015; 373:795-807

References

1. McArthur JC, Hoover DR, Bacellar H, et al. Dementia in AIDS patients: incidence and risk factors. Multicenter AIDS Cohort Study. *Neurology* 1993; **43**(11): 2245-52.
2. Portegies P, de Gans J, Lange JM, et al. Declining incidence of AIDS dementia complex after introduction of zidovudine treatment. *BMJ* 1989; **299**(6703): 819-21.
3. Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology* 2007; **69**(18): 1789-99.
4. Wilkie FL, Goodkin K, Eisdorfer C, et al. Mild cognitive impairment and risk of mortality in HIV-1 infection. *J Neuropsychiatry Clin Neurosci* 1998; **10**(2): 125-32.
5. Albert SM, Marder K, Dooneief G, et al. Neuropsychologic impairment in early HIV infection. A risk factor for work disability. *Arch Neurol* 1995; **52**(5): 525-30.
6. Hinkin CH, Castellon SA, Durvasula RS, et al. Medication adherence among HIV+ adults: effects of cognitive dysfunction and regimen complexity. *Neurology* 2002; **59**(12): 1944-50.
7. Tozzi V, Balestra P, Galgani S, et al. Positive and sustained effects of highly active antiretroviral therapy on HIV-1-associated neurocognitive impairment. *AIDS* 1999; **13**(14): 1889-97.
8. Robertson KR, Robertson WT, Ford S, et al. Highly active antiretroviral therapy improves neurocognitive functioning. *J Acquir Immune Defic Syndr* 2004; **36**(1): 562-6.
9. Cysique LA, Vaida F, Letendre S, et al. Dynamics of cognitive change in impaired HIV-positive patients initiating antiretroviral therapy. *Neurology* 2009; **73**(5): 342-8.
10. Cysique LA, Waters EK, Brew BJ. Central nervous system antiretroviral efficacy in HIV infection: a qualitative and quantitative review and implications for future research. *BMC Neurol* 2011; **11**: 148.
11. Ellis RJ, Letendre S, Vaida F, et al. Randomized trial of central nervous system-targeted antiretrovirals for HIV-associated neurocognitive disorder. *Clin Infect Dis* 2014; **58**(7): 1015-22.
12. Hellmuth J, Fletcher JL, Valcour V, et al. Neurologic signs and symptoms frequently manifest in acute HIV infection. *Neurology* 2016; **87**(2): 148-54.
13. Peluso MJ, Meyerhoff DJ, Price RW, et al. Cerebrospinal fluid and neuroimaging biomarker abnormalities suggest early neurological injury in a subset of individuals during primary HIV infection. *J Infect Dis* 2013; **207**(11): 1703-12.

14. Underwood J, Robertson KR, Winston A. Could antiretroviral neurotoxicity play a role in the pathogenesis of cognitive impairment in treated HIV disease? *AIDS* 2015; **29**(3): 253-61.
15. Group ISS, Lundgren JD, Babiker AG, et al. Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *N Engl J Med* 2015; **373**(9): 795-807.
16. Wright EJ, Grund B, Cysique LA, et al. Factors associated with neurocognitive test performance at baseline: a substudy of the INSIGHT Strategic Timing of AntiRetroviral Treatment (START) trial. *HIV Med* 2015; **16 Suppl 1**: 97-108.
17. Wright E, Brew B, Arayawichanont A, et al. Neurologic disorders are prevalent in HIV-positive outpatients in the Asia-Pacific region. *Neurology* 2008; **71**(1): 50-6.
18. Radloff L. The CES-D Scale: a self-report depression scalefor research in the general population. *Applied psychological measures* 1977; **1**: 385-401.
19. Carey CL, Woods SP, Gonzalez R, et al. Predictive validity of global deficit scores in detecting neuropsychological impairment in HIV infection. *J Clin Exp Neuropsychol* 2004; **26**(3): 307-19.
20. Letendre S. Central nervous system complications in HIV disease: HIV-associated neurocognitive disorder. *Top Antivir Med* 2011; **19**(4): 137-42.
21. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J* 1991; **121**(1 Pt 2): 293-8.
22. Team. RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. https://www.R-project.org/.
23. McCaffrey R, Duff K, Westervelt HJ. Practitioner's guide to evaluating change with neuropsychological assessment instruments. New York; 2000.
24. Heilbronner RL, Sweet JJ, Attix DK, Krull KR, Henry GK, Hart RP. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *Clin Neuropsychol* 2010; **24**(8): 1267-78.
25. Grund B, Wright EJ, Brew BJ, et al. Improved neurocognitive test performance in both arms of the SMART study: impact of practice effect. *J Neurovirol* 2013; **19**(4): 383-92.
26. Cysique LA, Franklin D, Jr., Abramson I, et al. Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *J Clin Exp Neuropsychol* 2011; **33**(5): 505-22.

27. Magnus Gisslén, Richard W. Price, Serena Spudich, Lars Hagberg, Steven G. Deeks, Peter Burbelo. HIV antibodies in CSF and serum in untreated and treated infection. CROI, February 13–16, 2017, Seattle, Washington, Abstract #391.

28. Gisslén M, Price RW, Andreasson U, Norgren N, Nilsson S, Hagberg L, Fuchs D, Spudich S, Blennow K, Zetterberg H. Plasma Concentration of the Neurofilament Light Protein (NFL) is a Biomarker of CNS Injury in HIV Infection: A Cross-Sectional Study. EBioMedicine. 2016;3:135-40.

29. Yilmaz A, Blennow K, Hagberg L, Nilsson S, Price RW, Schouten J, Spudich S, Underwood J, Zetterberg H, Gisslen M. Neurofilament light chain protein as a marker of neuronal injury: review of its use in HIV-1 infection and reference values for HIV-negative controls. Expert Rev Mol Diagn. 2017;17(8):761-70.

30. Van Zoest RA, Underwood J, De Francesco D, Sabin CA, Cole JH, Wit FW, Caan MWA, Kootstra NA, Fuchs D, Zetterberg H, Majoie C, Portegies P, Winston A, Sharp DJ, Gisslen M, Reiss P, Co-morbidity in Relation to AC. Structural brain abnormalities in successfully treated HIV infection: associations with disease and cerebrospinal fluid biomarkers. J Infect Dis. 2017. doi: 10.1093/infdis/jix553. PubMed PMID: 29069436.

31. Cole JH, Underwood J, Caan MW, De Francesco D, van Zoest RA, Leech R, Wit FW, Portegies P, Geurtsen GJ, Schmand BA, Schim van der Loeff MF, Franceschi C, Sabin CA, Majoie CB, Winston A, Reiss P, Sharp DJ, collaboration C. Increased brain-predicted aging in treated HIV disease. Neurology. 2017;88(14):1349-57.

32. Haddow LJ, Laverick R, Daskalopoulou M, McDonnell J, Lampe FC, Gilson R, Speakman A, Antinori A, Balestra P, Bruun T, Gerstoft J, Nielsen L, Vassilenko A, Collins S, Rodger AJ, Cognitive Impairment in People with HIV. Multicenter European Prevalence Study of Neurocognitive Impairment and Associated Factors in HIV Positive Patients. AIDS Behav. 2017. doi: 10.1007/s10461-017-1683-z. PubMed PMID: 28144792.

33. Heaton RK, Franklin DR, Jr., Deutsch R, Letendre S, Ellis RJ, Casaletto K, Marquine MJ, Woods SP, Vaida F, Atkinson JH, Marcotte TD, McCutchan JA, Collier AC, Marra CM, Clifford DB, Gelman BB, Sacktor N, Morgello S, Simpson DM, Abramson I, Gamst AC, Fennema-Notestine C, Smith DM, Grant I, Group C. Neurocognitive change in the era of HIV combination antiretroviral therapy: the longitudinal CHARTER study. Clin Infect Dis. 2015;60(3):473-80.

34. Moore DJ, Roediger MJ, Eberly LE, et al. Identification of an abbreviated test battery for detection of HIV-associated neurocognitive impairment in an early-managed HIV-infected cohort. *PLoS One* 2012; **7**(11): e47310.
35. WHO. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. WHO Press; 2015.

Figure 1. Study design and CONSORT flow diagram.

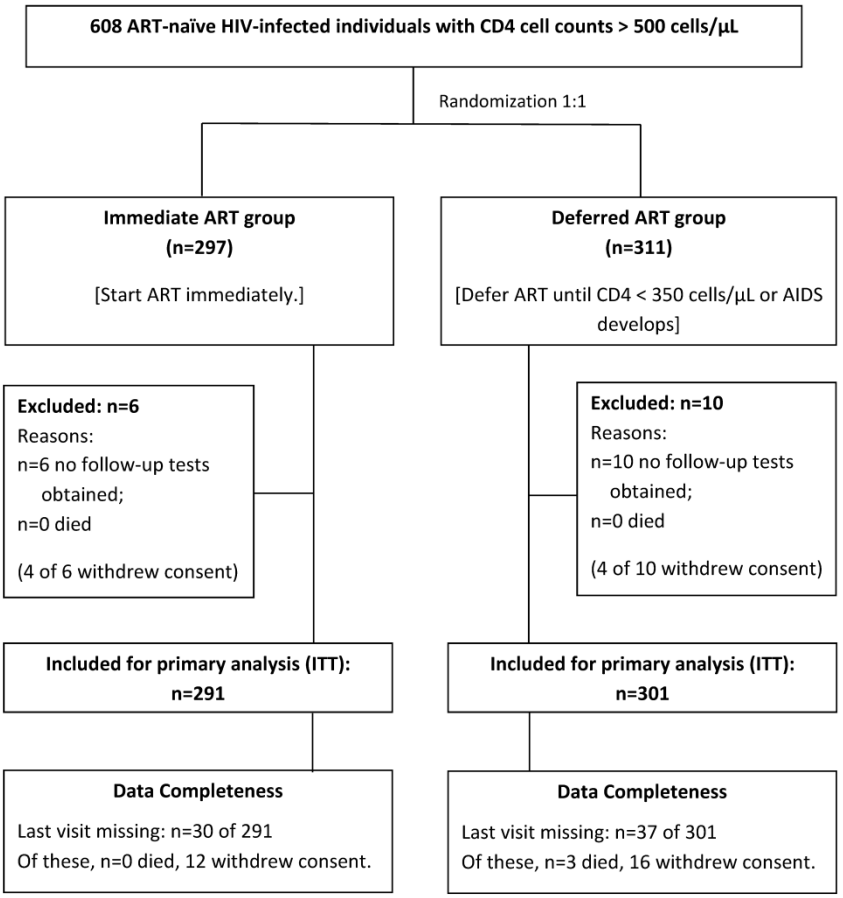
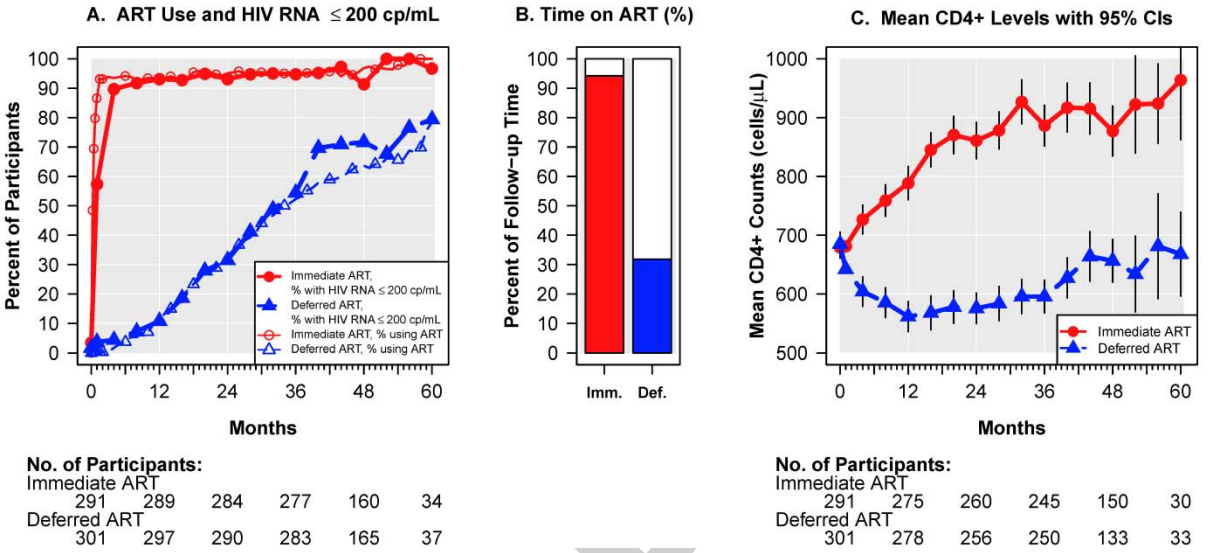


Figure 2. (A) Percent of participants using ART, and percent with suppressed viral load (HIV RNA $\leq$ 200 copies/mL) over time; (B) ART use expressed as percent of follow-up time accrued; (C) Mean CD4 cell count levels ($\pm$ 2 SE) over time.



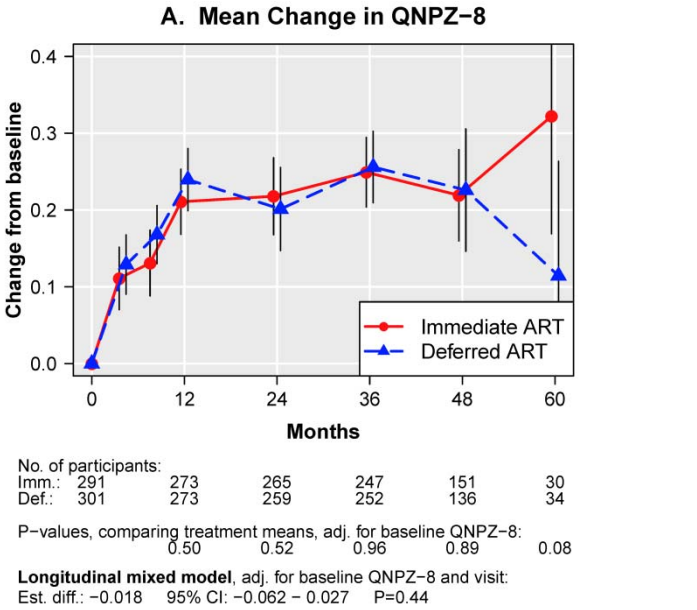
No. of Participants:

Immediate ART	291	289	284	277	160	34
Deferred ART	301	297	290	283	165	37

No. of Participants:

Immediate ART	291	275	260	245	150	30
Deferred ART	301	278	256	250	133	33

Figure 3.(A) Change in mean QNPZ-8 scores from baseline through follow-up; (B)Change in mean z-scores for the individual tests, which are averaged to calculate the QNPZ-8 summary score.



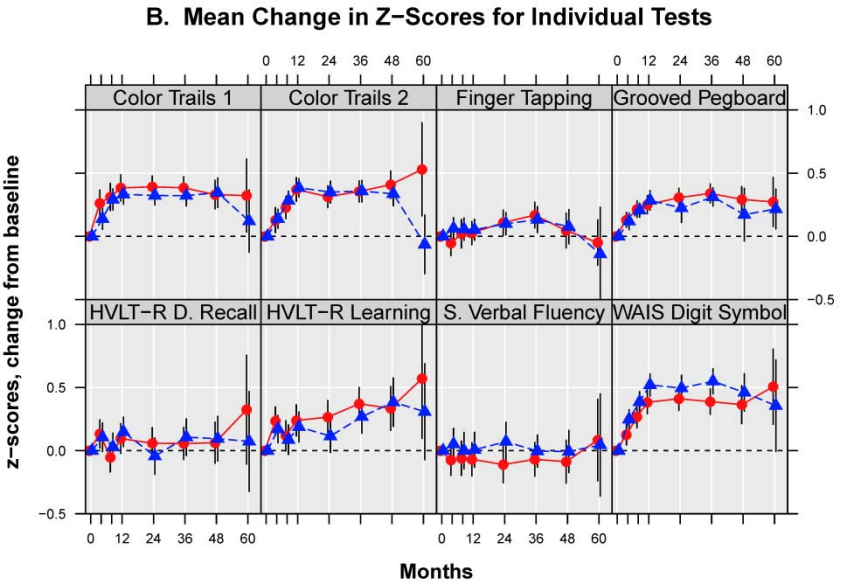


Figure 4. Subgroup analyses for change in mean QNPZ-8 scores from baseline.

Footnote to Fig. 4: When adjusting interaction p-values for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) method, $p \leq 0.004$ provides evidence for heterogeneity of the treatment effect across subgroups at the FDR$\leq$0.05 level. Subgroup analyses by age, education, HIV RNA level, pre-specified ART regimens, and their CNS penetration effectiveness score were specified a priori in the study protocol. In addition to the 8 subgroup factors shown, we analyzed subgroups by 16 baseline factors: by race,sex, employment status, urban residence, country of enrollment, time since HIV diagnosis, CD4 cell count, body mass index, diabetes, depression (CES-D$\geq$16), prior psychiatric diagnosis, prior cardiovascular disease, 10-year Framingham risk of CHD, hematocrit, AST/SGOT, ALT/SGPT.The treatment effect was homogeneous across those 16 subgroups.

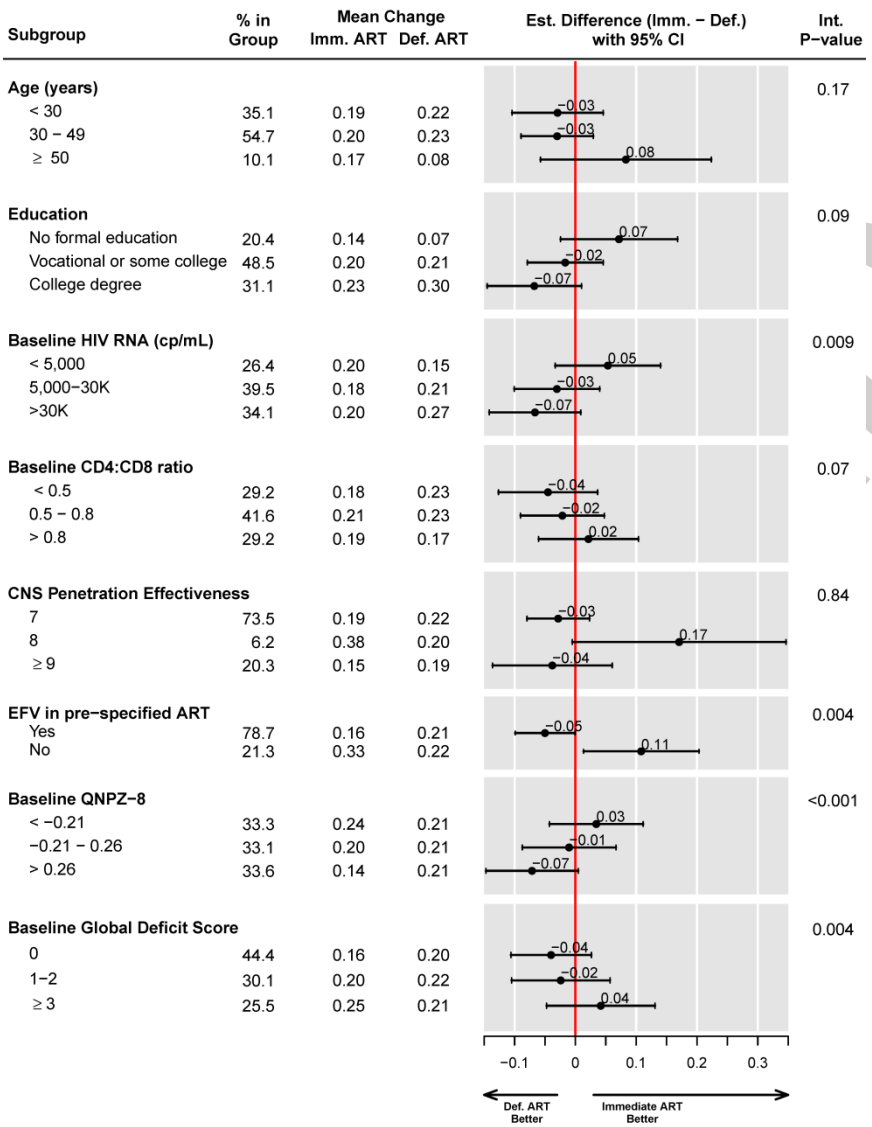


Table 1. Baseline Characteristics

Characteristic	Median [IQR] or N (%)		
	Immediate ART (n= 291)	**Deferred ART (n= 301)**	**Total (n= 592)**
Age (years)	33 [27 - 42]	35 [28 - 44]	34 [27 - 42]
Female (%)	27 (9.3%)	40 (13.3%)	67 (11.3%)
Race (%)			
Black	38 (13.1%)	52 (17.3%)	90 (15.2%)
Latino/Hispanic	49 (16.8%)	47 (15.6%)	96 (16.2%)
Asian	45 (15.5%)	50 (16.6%)	95 (16.0%)
White	140 (48.1%)	140 (46.5%)	280 (47.3%)
Other	19 (6.5%)	12 (4.0%)	31 (5.2%)
Highest formal training (%)			
No formal training	58 (19.9%)	63 (20.9%)	121 (20.4%)
Vocational training, completed	72 (24.7%)	73 (24.3%)	145 (24.5%)
Some college or university	73 (25.1%)	69 (22.9%)	142 (24.0%)
Bachelor's degree or equivalent	65 (22.3%)	71 (23.6%)	136 (23.0%)
Master's degree or higher	23 (7.9%)	25 (8.3%)	48 (8.1%)
Currently employed (%)	231 (79.4%)	221 (73.4%)	452 (76.4%)
Urban residence (%)	253 (86.9%)	266 (88.4%)	519 (87.7%)
Country of enrollment (%)			
United Kingdom/Australia	33 (11.3%)	33 (11.0%)	66 (11.1%)
European countries[1]	47 (16.2%)	55 (18.3%)	102 (17.2%)
United States	42 (14.4%)	45 (15.0%)	87 (14.7%)
Thailand	42 (14.4%)	47 (15.6%)	89 (15.0%)
Brazil	85 (29.2%)	80 (26.6%)	165 (27.9%)
Argentina/Chile	42 (14.4%)	41 (13.6%)	83 (14.0%)
Time since HIV diagnosis (years)	0.8 [0.2 - 2.5]	0.9 [0.3 - 2.5]	0.9 [0.3 - 2.5]
Likely mode of HIV infection (%)			
Injection drug use	4 (1.4%)	1 (0.3%)	5 (0.8%)
Male sexual contact, same sex	222 (76.3%)	220 (73.1%)	442 (74.7%)

Sexual contact, opposite sex	53 (18.2%)	62 (20.6%)	115 (19.4%)
Other/unknown	12 (4.1%)	18 (6.0%)	30 (5.1%)
CD4 (cells/μL)	632 [578 - 745]	628 [570 - 735]	629 [575 - 741]
Nadir CD4 (cells/ μL)	535 [466 - 626]	534 [473 - 638]	535 [471 - 631]
CD4:CD8 ratio	0.64 [0.46 - 0.84]	0.63 [0.47 - 0.85]	0.64 [0.47 - 0.84]
HIV RNA (copies/mL)	18126 [5260 - 46700]	13317 [3609 - 41357]	15441 [4595 - 44700]
Body mass index (kg/m^2)	23.8 [21.5 - 26.7]	23.7 [21.8 - 27.0]	23.8 [21.6 - 26.8]
Prior CVD diagnosis[2] (%)	5 (1.7%)	3 (1.0%)	8 (1.4%)
Framingham 10-year risk of CHD	1.7 [0.5 - 4.6]	2.0 [0.6 - 4.9]	1.8 [0.5 - 4.7]
Hepatitis B or C (%)	21 (7.3%)	16 (5.3%)	37 (6.3%)
Alcoholism/other substance dependence (%)	15 (5.2%)	16 (5.3%)	31 (5.2%)
Psychiatric diagnosis[3] (%)	28 (9.6%)	21 (7.0%)	49 (8.3%)
CES-D score[4]	10 [5 - 17]	10 [5 - 19]	10 [5 - 18]
CES-D score[4] ≥ 16 (%)	86 (31.6%)	92 (31.9%)	178 (31.8%)
GDS[5]	1 [0 - 2]	1 [0 - 3]	1 [0 - 3]
Mild impairment[6] (%)	50 (17.2%)	67 (22.3%)	117 (19.8%)
Moderate impairment[6] (%)	7 (2.4%)	9 (3.0%)	16 (2.7%)
CPE score[7], pre-specified ART regimen	7 [7 - 8]	7 [7 - 8]	7 [7 - 8]

[1] Germany, Italy, Belgium, and Switzerland.

[2] History of myocardial infarction, stroke, or coronary revascularization.

[3] Major depression, bipolar disorder, schizophrenia, or other psychotic disorder.

[4] Center for Epidemiological Studies Depression scale, ≥16 denotes depression.

[5] Global Deficit Score, average of deficit scores over 8 tests, where the deficit score for an individual test is defined by its z-scores, 0 for z≥-1,
1 for -1 > z ≥-1.5, 2 for -1.5 > z ≥ -2.0, 3 for -2.0 > z ≥ -2.5, 4 for -2.5 > z ≥ -3.0, 5 for z <-3.0.

[6]Z-scores below -1 (for mild impairment) or below -2 (for moderate impairment) for 2 or more of the 6 tested domains.

[7]Central nervous system penetration efficacy score